

# Multiple recurrences and the associated matrix structures stemming from normal matrices

Clara Mertens, and Raf Vandebril

## Abstract

There are many classical results in which orthogonal vectors stemming from Krylov subspaces are linked to short recurrence relations, e.g., three term recurrences for Hermitian and short rational recurrences for unitary matrices. These recurrence coefficients can be captured in a Hessenberg matrix, whose structure reflects the relation between the spectrum of the original matrix and the recurrences. The easier the recurrences, the faster the orthogonal vectors can be computed possibly resulting in computational savings in the design of, e.g., iterative solvers.

In this article we focus on multiple recurrence relations, i.e., the  $(j + 1)$ st orthogonal vector satisfies

$$\mathbf{q}_{j+1} = \sum_{i=j-m}^j \rho_{j,i} A \mathbf{q}_i - \sum_{i=j-\ell}^j \gamma_{j,i} \mathbf{q}_i,$$

with  $\rho_{j,i}, \gamma_{j,i}$  scalars, and  $A$  the matrix defining the Krylov space. Though many compelling results are around, the structure of the corresponding Hessenberg matrix is mostly deduced by analyzing the inner product relations.

In this article we first review classical results on short multiple recurrences for normal matrices whose Hermitian conjugate can be written as a ‘low degree’ rational function of the matrix. Instead of considering inner product relations, we reformulate this theory in a matrix setting. Moreover, the matrix building blocks allow us to also derive multiple recurrence relations for  $B$ -normal matrices, normal matrices whose eigenvalues lie on the union of curves in the complex plane, normal matrices perturbed by a low rank, and normal matrices  $A$  satisfying a ‘low degree’ relation  $s(A^*) = p(A)q(A)^{-1}$ .

The theoretical results on the structure available in the Hessenberg matrix lead to a new way to compute the orthogonal vectors. The numerical experiments illustrate, however, that straightforward algorithms exploiting this structure are numerically very sensitive, and more research is required to develop robust algorithms.

## Article information

- Mertens, Clara; Vandebril, Raf. Multiple recurrences and the associated matrix structures stemming from normal matrices, SIAM Journal on Numerical Analysis, volume 52, issue 6, pages 2977-2999, 2014.
- This article equals the final publisher’s version. A link is found below.
- Journal’s homepage: <https://www.siam.org/journals/sinum.php>
- Published version: <http://dx.doi.org/10.1137/120903555>
- KU Leuven’s repository url: <https://lirias.kuleuven.be/handle/123456789/486694>

## MULTIPLE RECURRENCES AND THE ASSOCIATED MATRIX STRUCTURES STEMMING FROM NORMAL MATRICES\*

CLARA MERTENS<sup>†</sup> AND RAF VANDEBRIL<sup>†</sup>

**Abstract.** There are many classical results in which orthogonal vectors stemming from Krylov subspaces are linked to short recurrence relations, e.g., three term recurrences for Hermitian and short rational recurrences for unitary matrices. These recurrence coefficients can be captured in a Hessenberg matrix, whose structure reflects the relation between the spectrum of the original matrix and the recurrences. The easier the recurrences, the faster the orthogonal vectors can be computed possibly resulting in computational savings in the design of, e.g., iterative solvers. In this article we focus on multiple recurrence relations, i.e., the  $(j + 1)$ st orthogonal vector satisfies  $\mathbf{q}_{j+1} = \sum_{i=j-m}^j \rho_{j,i} A \mathbf{q}_i - \sum_{i=j-\ell}^j \gamma_{j,i} \mathbf{q}_i$  with  $\rho_{j,i}$ ,  $\gamma_{j,i}$  scalars and  $A$  the matrix defining the Krylov space. Though many compelling results are around, the structure of the corresponding Hessenberg matrix is mostly deduced by analyzing the inner product relations. In this article we first review classical results on short multiple recurrences for normal matrices whose Hermitian conjugate can be written as a “low degree” rational function of the matrix. Instead of considering inner product relations, we reformulate this theory in a matrix setting. Moreover, the matrix building blocks allow us to also derive multiple recurrence relations for  $B$ -normal matrices, normal matrices whose eigenvalues lie on the union of curves in the complex plane, normal matrices perturbed by a low rank, and normal matrices  $A$  satisfying a “low degree” relation  $s(A^*) = p(A)q(A)^{-1}$ . The theoretical results on the structure available in the Hessenberg matrix lead to a new way to compute the orthogonal vectors. The numerical experiments illustrate, however, that straightforward algorithms exploiting this structure are numerically very sensitive, and more research is required to develop robust algorithms.

**Key words.** short multiple recurrences, Krylov subspaces, Hermitian, unitary, structured (rank) matrices

**AMS subject classifications.** 65F10, 65F25

**DOI.** 10.1137/120903555

**1. Introduction.** It is a tendency that matrices used in formulating solutions to particular applications grow bigger and bigger. A standard approach to deal with these often immense matrices is to project them onto matrices of manageable size. The most widespread subspaces used to project onto are the Krylov subspaces [24, 25, 27], leading to a variety of iterative methods, for both solving systems of equations and retrieving eigenvalues.

Some numerical issues, due to the involvement of the power method, cause the Krylov vectors (spanning the space) to be unattractive as basis vectors for creating the projection operator. Instead, a sequence of orthogonal vectors, iteratively generated, is constructed spanning the corresponding Krylov space and resulting in a

---

\*Received by the editors December 26, 2012; accepted for publication (in revised form) September 29, 2014; published electronically December 11, 2014. The research of the authors was partially supported by the Research Council KU Leuven, projects OT/11/055 (Spectral Properties of Perturbed Normal Matrices and Their Applications), CREA/13/012 (Can Unconventional Eigenvalue Algorithms Supersede the State-of-the-Art), PFV/10/002 (Optimization in Engineering, OPTEC); by the Fund for Scientific Research–Flanders (Belgium) project G034212N (Reestablishing Smoothness for Matrix Manifold Optimization via Resolution of Singularities); and by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office, Belgian Network DYSCO (Dynamical Systems, Control, and Optimization).

<http://www.siam.org/journals/sinum/52-6/90355.html>

<sup>†</sup>Department of Computer Science, KU Leuven, 3000 Leuven, Belgium (clara.mertens@cs.kuleuven.be, raf.vandebril@cs.kuleuven.be).

numerical more reliable projection operator. The simple (modified) Arnoldi method (with orthogonalization) is used for constructing these orthogonal vectors [2, 12, 25]. Unfortunately the computational complexity of Arnoldi increases rapidly as the iterative process proceeds. To remedy this, research has focused quite a lot on attempting to reduce the complexity of the Arnoldi method by exploiting additional matrix information.

Storing the recurrence coefficients in a matrix, let's call it the *recurrence* matrix, results in a Hessenberg structure. Being able to find “hidden” relations in the dense part of this matrix can give rise to significant speedup. For instance, it is well known nowadays that (almost) Hermitian [9, 18] matrices give rise to three term recurrences, resulting in a reduction from quadratic to linear complexity. Also for unitary matrices manners for generating the orthogonal vectors economically were proposed [6, 31]. Whereas the Hessenberg recurrence matrix becomes tridiagonal in the Hermitian case, it is typically dense in the unitary case. However, the upper triangular part of the unitary Hessenberg matrix is of structured rank form: all columns, up to the diagonal, are multiples of each other, allowing a *multiple recurrence*.

The search for structure and cheap algorithms led to many publications on this subject. The most well-known article is by Faber and Manteuffel [9]. Their analysis characterizes a subset of normal matrices allowing the construction of short recurrences for a CG-like algorithm [17]. Not only the sufficiency but also the necessity of matching the matrix structure—banded in this case—for inducing a particular type of recurrence is proved. Originally one focused mostly on finding sparse recurrence matrices [7, 8, 9, 18]. The existence of a low rank structure in the Hessenberg matrix [4, 5], however, allowed the design of *multiple recurrences*, where the  $(j + 1)$ st orthogonal vector satisfies

$$\mathbf{q}_{j+1} = \sum_{i=j-m}^j \rho_{j,i} A \mathbf{q}_i - \sum_{i=j-\ell}^j \gamma_{j,i} \mathbf{q}_i$$

with  $\rho_{j,i}$ ,  $\gamma_{j,i}$  scalars and  $A$  the matrix defining the Krylov space. These multiple recurrences are studied in [2] based on inner product relations.

In this article we will provide new insights into the structure of the matrix capturing the recurrence coefficients. Whereas most articles utilize inner products and orthogonality relations among the vectors, we base our deductions solely on “simple” matrix operations. We first review the classical results from Faber and Manteuffel. Based on the central theorem in this article, Theorem 3.1, the available low rank structure in the Hessenberg matrix is characterized precisely. Exploiting this structure not only results in multiple recurrences—identical to the ones of Faber and Manteuffel—but also provides a new algorithm to compute the orthogonal vectors. This new algorithm generalizes former work by Bella et al. [4, 5] by relaxing the extra *well-freeness* conditions as imposed in their article.

The matrix framework admits extending the applicability of the algorithm to several other cases:  $B$ -normal matrices, normal matrices whose eigenvalues lie on the union of curves in the plane, normal matrices perturbed by a low rank, and normal matrices  $A$  satisfying a “low degree” relation  $s(A^*) = p(A)q(A)^{-1}$  are discussed; corresponding recurrences are deduced; and the necessary adaptations on the algorithm are presented. In this context, the reader also might be interested in [10], where normal matrices are considered whose spectrum is contained in a second-degree curve.

The article is organized as follows. By a detailed examination in section 3 of the matrix structure resulting from multiplications, addition, and inversion we are able

to provide precise predictions of the structure of the recurrence matrix in section 4 whenever the Hermitian conjugate of the original matrix can be written as a rational function in the matrix itself. Relying on these accurate predictions of the structure we will derive the multiple recurrences and an algorithm to compute these orthogonal vectors in section 5. In section 6 we provide some more insight on particular cases; the structure of the recurrence matrix is deduced for matrices whose spectrum is decomposed into two parts, both satisfying a particular relation resulting in structure; low rank perturbations of normal matrices;  $B$ -normal matrices; and normal matrices satisfying a more general polynomial relation. Before concluding, some numerical experiments in section 7, serving as proof of concept, but also revealing numerical issues, are presented.

**2. Preliminaries and problem formulation.** The following notation is used throughout the article. Matrices are depicted by uppercase letters  $A = (a_{ij})$ , and  $I$  denotes the identity matrix. Vectors are boldface, lowercase letters, e.g.,  $\mathbf{x}, \mathbf{y}$ , and  $\mathbf{z}$ . With  $\text{span}\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$  the subspace generated by  $\mathbf{x}, \mathbf{y}, \mathbf{z}$  is meant. The standard inner product is denoted as  $\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{y}^* \mathbf{x}$  with  $\cdot^*$  the Hermitian conjugate. Sub- and superdiagonals frequently appear throughout the article. The elements  $h_{i,i+s}$  form the  $s$ th diagonal for positive  $s$ , also named the  $s$ th superdiagonal, and the  $(-s)$ th subdiagonal for negative  $s$ . The words sub- and superdiagonal are shorthand for the first sub- and superdiagonal, respectively.

Let us elaborate more on the key objects in this article: normal matrices, Krylov spaces, and recurrences. Polynomial relations between a matrix, its Hermitian conjugate, and its inverse are common when working with normal matrices. The typical case that will be studied in this article is  $A^* = p(A)q(A)^{-1}$  with  $p(\cdot)$  and  $q(\cdot)$  polynomials and  $q(A)$  invertible, e.g., a unitary matrix obeys  $A^* = A^{-1}$ , a Hermitian matrix satisfies  $A^* = A$ , and for a skew-Hermitian matrix we have  $A^* = -A$ . In this article classes of normal matrices will be characterized by these relations. We will, however, not search for the minimal rational function in any particular sense [19, 21]; we plainly assume the existence of such relations and establish our results thereon.

For  $A \in \mathbb{C}^{n \times n}$ ,  $\mathbf{v} \in \mathbb{C}^n$  we let

$$(2.1) \quad \mathcal{K}_k(A, \mathbf{v}) = \text{span}\{\mathbf{v}, A\mathbf{v}, A^2\mathbf{v}, \dots, A^{k-1}\mathbf{v}\}$$

denote the Krylov subspace of dimension  $k$  for  $A$  with initial vector  $\mathbf{v}$ . We also denote  $K_k = [\mathbf{v}, A\mathbf{v}, \dots, A^{k-1}\mathbf{v}]$ . The supplement  $(A, \mathbf{v})$  for identifying the matrix  $A$  and initial vector  $\mathbf{v}$  will be omitted when clear from the context. For simplicity we assume throughout the manuscript that no breakdowns occur when building the Krylov subspaces; thus  $A$  is nonderogatory and the initial vector  $\mathbf{v}$  has full grade with respect to  $A$ . Note that this assumption does not limit the analysis; if  $\mathbf{v}$  were of grade  $m < n$  the Krylov procedure would break down after  $m$  iterations, implying that an invariant subspace  $\mathcal{L}$  of  $A$  was found. Then all reasoning below applies to the induced operator  $A|_{\mathcal{L}}$ .

The orthonormal vectors  $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_n$  are constructed by sequentially orthogonalizing the columns of  $K_k$ ; hence, for all  $k \leq n$ ,

$$\text{span}\{\mathbf{v}, A\mathbf{v}, \dots, A^{k-1}\mathbf{v}\} = \text{span}\{\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_k\}.$$

Throughout the manuscript we will rely on the classical Arnoldi procedure to generate the orthogonal vector sequences; the focus will not be on the accurate construction of these vectors, i.e., we will not bother the reader with modified Arnoldi or any kind

of orthogonality reinforcement such as reorthogonalization [12]. The Arnoldi method can algorithmically be described as follows:

1.  $\mathbf{q}_1 := \mathbf{v}/\|\mathbf{v}\|$ .
2. For  $j = 1, \dots, n-1$  do
  - $\mathbf{q}_{j+1} := A\mathbf{q}_j - \sum_{i=1}^j h_{ij}\mathbf{q}_i$ , with  $h_{ij} = \langle A\mathbf{q}_j, \mathbf{q}_i \rangle = \mathbf{q}_i^* A\mathbf{q}_j$
  - $h_{j+1,j} := \|\mathbf{q}_{j+1}\|$
  - $\mathbf{q}_{j+1} := \mathbf{q}_{j+1}/h_{j+1,j}$ .

In each iteration  $\mathbf{q}_{j+1}$  is constructed by orthogonalizing  $A\mathbf{q}_j$  against the previously computed vectors. This implies that  $A\mathbf{q}_j$  can be written as a combination of the  $(j+1)$ st orthogonal vectors:

$$(2.2) \quad A\mathbf{q}_j = \sum_{i=1}^{j+1} h_{ij}\mathbf{q}_i.$$

Taking, however, additional features of  $A$  into account, one might be able to deduce shorter recurrences. For instance, a Hermitian and a unitary matrix allow this.

Equation (2.2) represents a single step of the Arnoldi method. Collecting all such steps from 1 to  $n$  one obtains

$$(2.3) \quad AQ = QH \text{ with } Q = [\mathbf{q}_1, \dots, \mathbf{q}_n]$$

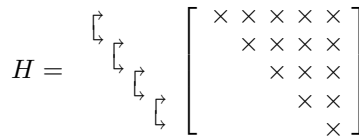
and  $H = (h_{ij})$  of Hessenberg form. The existence, however, of short recurrences is also reflected in the structure of the matrix  $H$ . It can become, e.g., of tridiagonal (e.g.,  $A$  Hermitian), banded, or structured rank (e.g.,  $A$  unitary) form.

The core contribution of this article is predicting in an alternative manner, relying on matrix language, the rank structure of the Hessenberg matrix  $H$ , and exploiting this rank structure to design an algorithm to retrieve short recurrences.

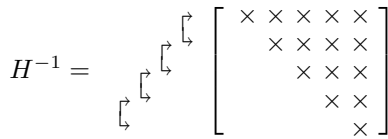
### 3. Building blocks for structure prediction of the Hessenberg matrix.

We will draw heavily from [29, 30]. The forthcoming proofs are mostly based on utilizing the  $QR$ -factorization of the associated matrices. To ease the understanding a graphical depiction is used. A Hessenberg matrix  $H$  admits a  $QR$ -factorization  $H = Q_1 \dots Q_{n-1}R$ , with each  $Q_i$  representing a  $2 \times 2$  (Givens) rotation acting on rows  $i$  and  $i+1$  embedded in the identity matrix. A matrix admitting a factorization of the form  $H^{-1} = \tilde{Q}_{n-1} \dots \tilde{Q}_1 \tilde{R}$  will be called an *inverse Hessenberg* matrix. Schematically both factorizations are depicted in Figure 1. Each bracket stands for a rotation, with arrowheads targeting the rows affected by the rotation.

The essence in the *detailed factorizations* of Figure 1 is the positioning of the individual rotations. The Hessenberg matrix has a *descending* sequence of rotations; the inverse Hessenberg matrix exhibits an *ascending* sequence of rotations. As breakdowns were not allowed when building the Krylov spaces, the associated Hessenberg



(a) Hessenberg.



(b) Inverse Hessenberg.

FIG. 1. Graphical depiction of two structured  $QR$ -factorizations.

matrices will be *unreduced* (or *irreducible* or *strict*), meaning that all subdiagonal elements differ from zero. This results in rotations different from the identity. Nonsingularity also implies  $R$  to be invertible and having thus nonzero diagonal elements.

The Hessenberg and inverse Hessenberg factorizations are linked via inversion. The missing link to prove this is the positioning: after inversion of a Hessenberg matrix the descending sequence becomes ascending but is located to the right of the upper triangular matrix. Rotations can, however, be *passed through* an upper triangular matrix, without altering the *pattern* of rotations. Consider, for instance,  $R$  upper triangular and a single nonidentity rotation  $Q_i$ . Explicitly forming the product  $RQ_i$  creates a matrix with a bulge in position  $(i+1, i)$ . This bulge can be removed by executing a nonidentity rotation on the left resulting in  $\tilde{Q}_i \tilde{R} = RQ_i$ . The matrices  $Q_i$  and  $\tilde{Q}_i$  act on the same rows. Passing rotations one by one through an upper triangular matrix, one can move an entire pattern of rotations from right to left or from left to right, preserving the pattern but altering the individual rotations.

**3.1. Structure transfer through inversion.** In this manuscript we will heavily rely on the transport of matrix structures through inversion. The nullity theorem, formulated in its matrix form by Fiedler and Markham [11], is one of the most simple yet most powerful theorems in this context. In the following lemma, a direct consequence of the nullity theorem is proposed.  $|\cdot|$  reflects the cardinality of a set, and by  $A(\alpha, \beta)$  standard submatrix selection based on index sets  $\alpha$  and  $\beta$  is meant.

LEMMA 3.1. *For an invertible matrix  $A$  of size  $n \times n$*

$$\text{rank}(A^{-1}(\alpha, \beta)) = \text{rank}(A(N \setminus \beta, N \setminus \alpha)) + |\alpha| + |\beta| - n,$$

where  $\alpha$  and  $\beta$  depict a selection of indices from  $N = \{1, \dots, n\}$ .

The statements on inversion without proof in this subsection can be deduced in a straightforward manner making use of this theorem. A more elaborate overview of the interplay between ranks of matrices and their inverses can be found in [29, 26] and the references therein.

COROLLARY 1. *The inverse of a Hessenberg matrix has all submatrices taken out of the part below the first superdiagonal of rank at most one, i.e., it is an inverse Hessenberg matrix.*

The proof is straightforward by Lemma 3.1 or by examining the factorization in Figure 1(b). Nonetheless, Corollary 1 is insufficient for our purposes, we demand strict control over the positioning of the low rank blocks.

COROLLARY 2. *The inverse of an unreduced Hessenberg matrix satisfies  $(i = 1, \dots, n)$*

$$\text{rank}(A^{-1}(i : n, 1 : i)) = 1.$$

*This signifies that all blocks taken from the part strictly below the first superdiagonal and including the bottom left element  $A^{-1}(n, 1)$  have rank one.*

For instance, the unitary matrix

$$(3.1) \quad U = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad \text{with } U^{-1} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

illustrates that inclusion of the bottom left element of the inverse is essential to get strict rank constraints.

COROLLARY 3. *The inverse of a generalized Hessenberg matrix, having all matrix entries below the  $s$ th subdiagonal zero is a structured rank matrix, having all submatrices taken out of the part below the  $s$ th superdiagonal of rank at most  $s$ .*



FIG. 2. Graphical depiction of the generalized (inverse) Hessenberg factorization.

If the Hessenberg matrix were unreduced, i.e., its  $s$ th subdiagonal contains no zero elements, its inverse would satisfy ( $i = 1, \dots, n - s + 1$ )

$$\text{rank}(A^{-1}(i : n, 1 : i + s - 1)) = s,$$

or equivalently, all submatrices taken from the part below the  $s$ th superdiagonal containing the bottom left  $s \times s$  block  $A^{-1}(n - s + 1 : n, 1 : s)$  have rank  $s$ .

Again Lemma 3.1 yields this result. We will refer to a generalized Hessenberg matrix and its inverse as an  $s$ -(inverse) Hessenberg matrix. The unitary matrix  $[\mathbf{e}_3, \mathbf{e}_4, \mathbf{e}_5, \mathbf{e}_6, \mathbf{e}_1, \mathbf{e}_2]$  with  $\mathbf{e}_i$  the  $i$ th basis vector is an interesting example of Corollary 3. Considering, for instance,  $s = 3$ , the detailed factorization of an  $s$ -Hessenberg matrix in Figure 2(a), and using the same reasoning as in section 3 results in a the detailed factorization of an  $s$ -inverse Hessenberg matrix Figure 2(b) satisfying the desired rank constraints. Furthermore, one can infer from the figures that an  $s$ -Hessenberg matrix admits a detailed factorization with  $s$  descending sequences of rotations, each peeling off one subdiagonal. The  $s$ -inverse Hessenberg matrix, on the other hand, exhibits a pattern with three ascending sequences each tied to a rank one increase in the associated structured rank matrix.

Even though Figure 2 visualizes descending and ascending sequences of decreasing length, there is no loss of generality in considering them of full length. The constraints on the bandwidth and the rank remain identical.

**3.2. Structure transfer through rational matrix functions.** In this article we heavily build upon the structure of  $p(H)$  and  $p(H)q(H)^{-1}$ .

**LEMMA 3.2.** *For a polynomial  $p(\cdot)$  of degree  $s$  and  $H$  Hessenberg, the matrix  $p(H)$  has all elements below the  $s$ th subdiagonal zero; thus  $p(H)$  is  $s$ -Hessenberg.*

*For an unreduced Hessenberg  $H$ ,  $p(H)$  will also be unreduced.*

Explicitly computing  $p(H)$  immediately provides the result. Alternatively, considering the product of  $s$  Hessenberg matrices in detailed format, one can pass through all rotations to the outer left, resulting in a detailed factorization of an  $s$ -Hessenberg matrix. This detailed factorization will be composed of  $s$  descending sequences, each creating an extra diagonal when applied to the associated upper triangular matrix.

The following theorem generalizes Corollary 3 and determines the rank structure for a quotient of the form  $p(H)q(H)^{-1}$  with  $H$  Hessenberg. This will allow us to predict the rank structure of an arbitrary normal Hessenberg matrix satisfying a general relation of the form  $H^* = p(H)q(H)^{-1}$ . In the following section we will exploit this rank structure to retrieve short recurrences for calculating an orthonormal Krylov basis.



**THEOREM 3.1.** *Let  $H$  be a unreduced Hessenberg. Let  $p(\cdot)$  and  $q(\cdot)$  be coprime polynomials of degree  $\ell$  and  $m$ , respectively, such that  $q(H)$  is invertible. Define*

$$Z = p(H)q(H)^{-1}.$$

*Then each submatrix of the form  $Z(i : n, 1 : j)$  below the  $(m - \ell)$ th diagonal of  $Z$  such that  $\min(n - i + 1, j) \geq m$ , i.e., it contains the bottom left  $m \times m$  block of  $Z$ , has rank precisely  $m$ .*

*Proof.* The proof is by induction on the sum of the degrees of  $p(\cdot)$  and  $q(\cdot)$ .

- *Case 1:*  $\ell + m = 1$ .

If  $\ell = 1$  and  $m = 0$ , the result follows by Lemma 3.2 above. Indeed,  $Z$  will have all elements below the first subdiagonal equal to zero, which is the degree of  $q(\cdot)$ . If  $\ell = 0$  and  $m = 1$ , Lemma 3.1 above provides us the equality

$$\text{rank}(q(H)^{-1}(i : n, 1 : j)) = \text{rank}(q(H)(j + 1 : n, 1 : i - 1)) + j - i + 1$$

for  $j \leq i$ . As the first subdiagonal of  $q(H)$  consists of nonzero elements, one checks easily that all blocks of the form  $q(H)(j + 1 : n, 1 : i - 1)$  have rank  $i - j$ . This is illustrated in Figure 3.

Hence, the desired result follows.

- *Case 2:*  $\ell + m > 1$ ; we discern two additional cases,  $\ell \geq m$  and  $m > \ell$ .

$\ell \geq m$ : Using the Euclidean division algorithm we can write

$$p(H)q(H)^{-1} = \tilde{p}(H) + r(H)q(H)^{-1},$$

where  $\tilde{p}(\cdot)$  has degree  $\ell - m$  and  $r(\cdot)$  is a polynomial with degree  $s$ , where  $s < m$ . The polynomial  $r(\cdot)$  is nonzero due to the fact that  $p(\cdot)$  and  $q(\cdot)$  are coprime. Using the induction hypothesis,  $r(H)q(H)^{-1}$  has the desired rank structure below the  $(m - s)$ th diagonal. By Lemma 3.2 above,  $\tilde{p}(H)$  is a generalized Hessenberg matrix with zeros below the  $(\ell - m)$ th subdiagonal. Moreover, since  $s$  is strictly smaller than  $m$ , the  $(m - s)$ th diagonal is situated in the upper triangular part of the matrix, strictly above the  $(\ell - m)$ th subdiagonal. Therefore, the total sum  $p(H)q(H)^{-1}$  will have the desired rank structure below the  $(m - \ell)$ th diagonal.

$m > \ell$ : Our aim is to prove that all submatrices of the form  $p(H)q(H)^{-1}(i : n, 1 : j)$  with  $j \leq i + m - \ell - 1$  and  $\min(n - i + 1, j) \geq m$ , i.e., all submatrices below the  $(m - \ell)$ th diagonal containing the bottom left  $m \times m$  block of  $p(H)q(H)^{-1}$ , have rank  $m$ . Applying Lemma 3.1 above we obtain

$$(3.2) \quad \begin{aligned} \text{rank}(p(H)q(H)^{-1}(i : n, 1 : j)) \\ = \text{rank}(q(H)p(H)^{-1}(j + 1 : n, 1 : i - 1)) + j - i + 1. \end{aligned}$$

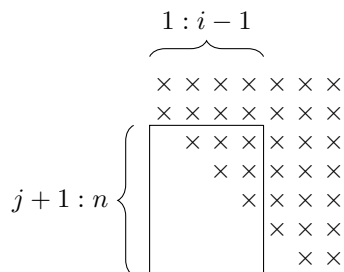
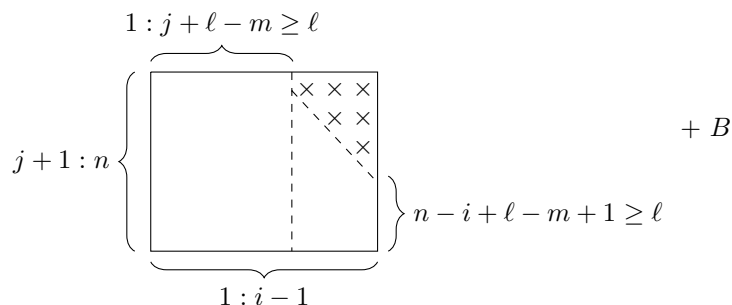


FIG. 3. Illustration of the nullity theorem.



FIG. 4. Structure of the submatrix  $q(H)p(H)^{-1}(j+1:n, 1:i-1)$ .

Therefore, in order to complete the proof one has to determine the ranks of the submatrices  $q(H)p(H)^{-1}(j+1:n, 1:i-1)$ , where  $i, j$  satisfy the inequalities  $j \leq i+m-\ell-1$ ,  $m \leq n-i+1$ , and  $m \leq j$ . In Figure 4 such a submatrix is depicted. Above the dashed diagonal corresponds to being above and including the  $(m-\ell)$ th subdiagonal.

By the Euclidean division algorithm we can write

$$(3.3) \quad q(H)p(H)^{-1} = \tilde{q}(H) + r(H)p(H)^{-1},$$

where  $\tilde{q}(\cdot)$  has degree  $m-\ell$  and  $r(\cdot)$  has degree  $s$  with  $s < \ell$ . By the induction hypothesis one knows that all submatrices of  $r(H)p(H)^{-1}$  below the  $(\ell-s)$ th diagonal containing the bottom left  $\ell \times \ell$  block have rank precisely  $\ell$ . Also,  $\tilde{q}(H)$  is an unreduced Hessenberg with zeros below the  $(m-\ell)$ th subdiagonal. As the  $(\ell-s)$ th diagonal is located strictly above the  $(m-\ell)$ th subdiagonal, (3.3) implies that  $q(H)p(H)^{-1}(j+1:n, 1:i-1)$  can be written as the sum of a matrix  $B$  and an invertible upper triangular matrix in the upper right corner. The matrix  $B$  is of the same size as  $q(H)p(H)^{-1}(j+1:n, 1:i-1)$  and has rank  $\ell$ , and its  $\ell \times \ell$  subblock in the lower left corner is nonsingular. This is visualized in Figure 4.

Next we claim that the rank of the matrix  $q(H)p(H)^{-1}(j+1:n, 1:j+\ell-m+k-1)$  is equal to  $\ell+k-1$ . We prove this by induction on  $k$ . For  $k=1$  this is clearly true. Assume that the matrix  $q(H)p(H)^{-1}(j+1:n, 1:j+\ell-m+k-1)$  has rank  $\ell+k-1$ . Suppose the rank of the matrix  $q(H)p(H)^{-1}(j+1:n, 1:j+\ell-m+k)$  also equals  $\ell+k-1$ . This implies that the columns of  $q(H)p(H)^{-1}(j+1:n, 1:\ell)$  span a vector of the form

$$\underbrace{(x_1, x_2, \dots, x_{k-1}, 1, 0, \dots, 0)}_{\text{length } n-j}.$$

Hence, there exists a nonzero linear combination of the columns of the submatrix  $q(H)p(H)^{-1}(k+1:n, 1:\ell)$  that equals the zero vector. This gives a contradiction, as these columns are linearly independent by the induction hypothesis. With the above observations, it follows that the rank of the matrix  $q(H)p(H)^{-1}(j+1:n, 1:i-1)$  is equal to  $\ell+(i-1-j-\ell+m) = -j+i-1+m$ . Substituting this into (3.2), we obtain

$$\text{rank}(p(H)q(H)^{-1}(i:n, 1:j)) = -j+i-1+m+j-i+1 = m,$$

which finishes the proof.  $\square$

**4. Krylov spaces and associated structure of the recurrence matrix.** In this section some known results for standard Krylov subspaces and recurrences are reconsidered, relying on the matrix structure only. We conclude with the general setting  $A^* = p(A)q(A)^{-1}$ .

**4.1.  $A^* = p(A)$  for a polynomial of degree  $\ell$ .** Every normal matrix satisfies  $A^* = p(A)$  for a polynomial of degree at most  $n - 1$ ; see, e.g., [16]. The relation  $A^* = p(A)$  passes on to the Hessenberg matrix  $H$ , hence  $H^* = p(H)$ . Degrees of  $p(\cdot)$  strictly lower than  $n - 1$  are reflected in the sparsity pattern of the upper triangular Hessenberg structure [22]. More precisely, by Lemma 3.2 and the relation  $H^* = p(H)$  we know that the upper bandwidth of  $H$  is restricted to  $\ell$ .

For example, consider the specific case where  $p(\cdot)$  is a polynomial of degree 1. It was shown in [9, 18, 23] that a matrix  $A$  with  $A^* = p(A)$  should be of the form  $A = \alpha I + e^{j\theta} S$ , with  $S$  Hermitian, or in other words  $A$  has collinear eigenvalues. This implies that  $H$  becomes of tridiagonal, not necessarily Hermitian, form, resulting in a three term recurrence.

**4.2. Unitary matrices.** Polynomial relations in  $A$  such as  $A^* = p(A)$  tell us in general something about the sparsity pattern of the Hessenberg matrix  $H$ . For a unitary matrix  $A$ , the degree of the polynomial  $p(\cdot)$  is  $n - 1$ ; however, writing  $A^*$  as a rational function of  $A$  provides us  $A^* = A^{-1}$ , which is clearly more elegant, and obviously needs fewer parameters to characterize  $A^*$ .

Unitarity of  $A$  transports to the Hessenberg matrix  $H$ . Hence  $H^* = H^{-1}$  and by Lemma 2 we know that the upper triangular part of the matrix  $H$  is of structured rank form: any subblock taken out of the upper triangular part of this matrix, including the diagonal, is of rank at most one. One can exploit the rank one structure of the Hessenberg  $H$  to compute a short recurrence of the form  $\mathbf{q}_{j+1} = \rho_{j-1} A \mathbf{q}_{j-1} + \rho_j A \mathbf{q}_j + \gamma_j q_j$ . The existence of such a recurrence is linked to the presence of the low rank part and will be proved in section 5. Short recurrences of this form, involving more than once the term  $A \mathbf{q}_i$ , are named *multiple recurrences* (see section 5 for a proper definition). In literature the existence of such a recurrence is already proven, relying on inner products and Krylov subspace properties; this in contrast to our approach, which is based on the matrix structure only.

**4.3. General case.** The following theorem summarizes all previously described cases and describes the rank structure of any normal Hessenberg matrix (see [22, 2]).

**THEOREM 4.1.** *Let  $A$  be a normal matrix satisfying  $A^* = p(A)q(A)^{-1}$  with  $p(\cdot)$  of degree  $\ell$  and  $q(\cdot)$  of degree  $m$ ,  $q(A)$  being invertible. Then the corresponding Hessenberg matrix  $H$  has all submatrices taken out of the part strictly above the  $(\ell - m)$ th diagonal of rank at most  $m$ .*

*Proof.* This is an immediate consequence of Theorem 3.1.  $\square$

It was also shown in [2], where these matrices were named  $(\ell, m)$ -normal, that these matrices admit a short recurrence of the form (5.1). Founded on Theorem 4.1 we will construct these short recurrences in an alternative way, making use of the structure of the corresponding Hessenberg matrix.

Note that Theorem 3.1 implies that all submatrices lying above the  $(\ell - m)$ th diagonal containing the upper right  $m \times m$  subblock have rank precisely  $m$ . This implies that the last  $m$  columns of the Hessenberg matrix  $H$  “generate” the low rank structure, i.e., the part of each column lying above the  $(\ell - m)$ th diagonal is a linear combination of these columns. This relation enables the design of short recurrences.

**4.4. Example.** As an example we consider a normal matrix  $A$  of which all eigenvalues lie on a circle in the complex plane. This is, for example, the case with shifted and scaled unitary matrices, i.e., matrices of the form  $\rho U + \xi I$  with  $U$  unitary.

Suppose this circle has center  $m$  and radius  $r$ . Then all eigenvalues  $\lambda$  satisfy the equation  $(\lambda - m)(\bar{\lambda} - \bar{m}) = r^2$ , or equivalently  $\bar{\lambda} = r^2/(\lambda - m) + \bar{m}$ . As the previous relation reveals, the corresponding Hessenberg matrix will have a rank one structure above and including the first superdiagonal. All subblocks in the upper triangular part with lower left corner on the main diagonal will have rank two unless all eigenvalues lie on a circle with the origin as center ( $m = 0$ ).

As we will prove in the forthcoming section, this implies that one can retrieve an orthonormal Krylov basis with a short recurrence of the form  $\mathbf{q}_{i+1} = \rho_i A \mathbf{q}_i + \rho_{i-1} A \mathbf{q}_{i-1} + \gamma_i \mathbf{q}_i + \gamma_{i-1} \mathbf{q}_{i-1}$ , where  $\rho_i, \gamma_i$  are parameters to be determined.

**5. Retrieving multiple recurrence relations relying on the matrix structure.** The starting point of [2] is almost identical to the one of this article, namely, is based on a relation for normal matrices trying to derive short recurrences, called *multiple recurrences*. An  $(\ell, m)$ -multiple recurrence signifies that each new vector can be written as

$$(5.1) \quad \mathbf{q}_{j+1} = \sum_{i=j-m}^j \rho_{j,i} A \mathbf{q}_i - \sum_{i=j-\ell}^j \gamma_{j,i} \mathbf{q}_i$$

with  $\rho_{j,i}$  and  $\sigma_{j,i}$  scalars. The proof and construction of these recurrences as done in [2] heavily depended on the usage of inner product relations. In [4, 5] a related problem was tackled by searching for recurrence relations linked to quasi-separable matrices. Additional conditions (named *well-free*), equivalent to the ones we will impose in section 5.1, were necessary to retain efficiency. In section 5.2 we will refrain from imposing any constraint and deal with the general setting.

**5.1. Recurrences in absence of difficulties.** We first tackle the trouble-free case and illustrate the idea with the unitary Hessenberg case.

**5.1.1. Unitary matrices.** As a unitary Hessenberg matrix is of structured rank form, a short recurrence can easily be obtained. Suppose we want to retrieve the  $j$  column of the matrix  $H$ . This column is pointed at by the arrow in (5.2).

$$(5.2) \quad H = \begin{array}{ccccccc} & & & & \downarrow & & \\ & \times & \times & \boxed{\times} & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times \\ & \times & \times & \times & \times & \times & \times \\ & & \times & \times & \times & \times & \times \\ & & & \times & \times & \times & \times \\ & & & & \times & \times & \times \\ & & & & & \times & \times \\ & & & & & & \times \\ & & & & & & & \times \end{array}$$

As  $H$  is structured the framed submatrix  $H(1 : j-1, j-1 : j-1)$  has rank at most one. Let us assume for simplicity that the rank is one, i.e., using the terminology of [5] the matrix is well-free. Then there exists an  $\alpha$  such that

$$H(1 : j-1, j-1 : j) \begin{bmatrix} -\alpha \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

Using  $A\mathbf{q}_{j-1} = \sum_{i=1}^j h_{i,j-1}\mathbf{q}_i$ , we obtain

$$\begin{aligned} A\mathbf{q}_j &= \sum_{i=1}^{j+1} h_{ij}\mathbf{q}_i = \sum_{i=1}^{j-1} h_{ij}\mathbf{q}_i + h_{jj}\mathbf{q}_j + h_{j+1,j}\mathbf{q}_{j+1} \\ &= \alpha(A\mathbf{q}_{j-1} - h_{j,j-1}\mathbf{q}_j) + h_{jj}\mathbf{q}_j + h_{j+1,j}\mathbf{q}_{j+1}. \end{aligned}$$

Hence we have

$$h_{j+1,j}\mathbf{q}_{j+1} = -\alpha A\mathbf{q}_{j-1} + A\mathbf{q}_j + (\alpha h_{j,j-1} - h_{jj})\mathbf{q}_j,$$

which is a recurrence of the form (5.1) with  $\ell = 0$  and  $m = 1$ .

Notice that the existence of this short recurrence is based on the fact that we assumed the framed submatrices in (5.2) to have rank precisely one. If the rank turns out to be zero, this is called *rank-deficient* in subsection 5.2, and our reasoning above no longer holds. These cases are excluded in [5].

**5.1.2.  $A^* = p(A) q(A)^{-1}$  with  $p(\cdot)$  of degree  $\ell$  and  $q(\cdot)$  of degree  $m$ .** The Hessenberg matrix  $H$  has all submatrices taken out of the part above the  $(\ell - m)$ th diagonal of rank at most  $m$ . Using this information about the structure of the matrix  $H$ , we will construct a recurrence of the form (5.1). To illustrate the construction of the coefficients we consider an example with  $\ell = 0$ ,  $m = 2$ . All submatrices above the dashed subdiagonal in (5.3) have rank at most two. As before, assume for simplicity the rank equals two. Let us now compute  $\mathbf{q}_6$ . As the framed submatrix  $H(1 : 4, 3 : 4)$  has rank two there exist  $\alpha_1$  and  $\alpha_2$  such that

$$(5.3) \quad H(1 : 4, 3 : 5) \begin{bmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

$\downarrow$

Since  $A\mathbf{q}_5 = \sum_{i=1}^6 h_{i5}\mathbf{q}_i$ , we obtain

$$A\mathbf{q}_5 = \sum_{i=1}^4 h_{i5}\mathbf{q}_i + h_{55}\mathbf{q}_5 + h_{65}\mathbf{q}_6 = \alpha_1 A\mathbf{q}_3 + \alpha_2 (A\mathbf{q}_4 - h_{54}\mathbf{q}_5) + h_{55}\mathbf{q}_5 + h_{65}\mathbf{q}_6,$$

which is of the form (5.1). In general, if  $j > \ell + 1$ , we select the submatrix  $H(1 : j - \ell - 1, j - m : j - 1)$  of the Hessenberg matrix  $H$  having rank at most  $m$ . Then we search for coefficients  $\alpha_i$ , for  $i = j - m, \dots, j$ , such that

$$H(1 : j - \ell - 1, j - m : j) \begin{bmatrix} -\alpha_{j-m} \\ \vdots \\ -\alpha_j \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}.$$

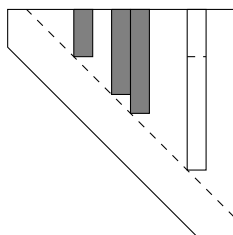


FIG. 5. An example of a Hessenberg matrix illustrating the algorithm.

Then one checks readily using the construction as illustrated before that

$$h_{j+1,j}\mathbf{q}_{j+1} = \sum_{i=j-m}^j \alpha_i A\mathbf{q}_i - \sum_{i=j-\ell}^j \gamma_i \mathbf{q}_i$$

for some  $\gamma_i$ , which can be determined by taking inner products with  $\mathbf{q}_j$  on both sides. For  $j \leq \ell + 1$  we can just use the classical Arnoldi recurrence, which is of the form (5.1) since  $j \leq \ell + 1$ .

**5.2. Dealing with rank deficiency.** Consider the matrix (3.1), which clearly exhibits troublesome behavior as all blocks needed for determining the next column are not of rank exactly one. In general a rank condition is imposed, stating that all blocks above a certain diagonal, including the upper right block of correct size, are of a fixed predetermined rank. Unfortunately, no conditions are put on submatrices not incorporating the upper right part. We will name such blocks with a rank strictly less than the expected rank *rank deficient*. For instance, (3.1) has several of these blocks. Encountering a rank deficient block during the execution of the algorithm in section 5.1 implies that from a certain point the new  $A\mathbf{q}_j$  will not necessarily be efficiently computable from the previous columns. We see, however, that for the matrix (3.1), the last column will *generate* the rank one structure, which means the part of each column located in the upper triangular part of the matrix will be a multiple of the upper part of the last column with corresponding length.

In this section a modification of the algorithm from section 5.1 is presented, which is able to cope with the problem of rank deficiencies. The aim of the algorithm is to fastly retrieve *generators* (vectors) for the low rank part of the considered Hessenberg matrix, such that successive columns can be written as linear combinations of these generators. The efficiency of the algorithm heavily depends on the number of columns needed to be computed before these generators are found. Unfortunately columns are computed one after the other, and if, as in (3.1), columns feasible as generators are only located in the trailing columns, efficiency can suffer dramatically. In [4, 5] these matrices are excluded and only so-called well-free matrices are discussed.

We now illustrate the idea of the algorithm. Consider a Hessenberg matrix of which the columns need to be efficiently computed in order to retrieve an orthonormal Krylov basis. Let us assume that the blocks above a specified diagonal should be of rank  $m$ ; this is indicated by the dashed diagonal line in Figure 5. Note that we only consider part of the vectors in the rank structured form; the remaining components are trivially computed using the Arnoldi method.

Our aim is to find  $m$  generators. A generator consists of the part of a column lying above the dashed diagonal, i.e., in the rank structured part of the Hessenberg matrix. Note that these generators have varying length. In Figure 5 three of these generators are indicated in gray. We also require these generators to be linearly independent, by which we mean they are linearly independent if we only take into account their first  $s$  components, where  $s$  denotes the length of the shortest generator. Once  $m$  generator columns are found, subsequent columns can be efficiently retrieved. Indeed, consider the white column indicated in Figure 5. Then clearly the first  $s$  components of this column can be written as a linear combination of the generators' first  $s$  components. To retrieve the linear dependencies an  $m \times m$  system of equations has to be solved. To solve this system some elements need to be computed explicitly, whereas the remaining components among the first  $s$  ones are retrieved via the systems' solution. All components below the  $s$ th element are computed using Arnoldi, which emphasizes that for computational efficiency it is important to have generators of maximal length. Therefore, once we have computed a new column, we have to decide whether it is possible to update the set of generators.

We now give a more rigorous description of the algorithm.

*Step 1.* Use Arnoldi to compute the first  $\ell$  columns, since these columns' structured rank parts have length less than  $m$ , which makes them unusable as generators.

*Step 2.* Next the algorithm searches for  $m$  generators. Let us assume that so far we have found  $k$  generators with  $k \leq m$ . Compute the next column of the Hessenberg matrix and check whether its first  $s$  components are a linear combination of the generators already found. The variable  $s$  is still used to denote the length of the shortest generator. If this is not the case, a new generator has been found and we go from  $k$  to  $k + 1$  generators. However, if this new column's first  $s$  components are a unique nonzero linear combination of the generators, then select a generator which contributes sufficiently enough in this linear combination, preferably with smallest possible length, and replace it by the column just computed. Consequently, we have updated our set of generators with the purpose of finding generators with maximal length.

*Step 3.* Once  $m$  generators are found, subsequent columns can be efficiently computed. To compute a new column, compute  $m$  elements using Arnoldi. Next an  $m \times m$  linear system has to be solved to write the new column as a linear combination of the generators. In practice, one can choose to solve an  $\tilde{m} \times m$  least squares problem with  $\tilde{m} \geq m$  to improve accuracy. Given the solution of this system of equations the remaining components up to and including the  $s$ th component of the new column are retrieved as the corresponding linear combination of the generators. The remaining components are computed using Arnoldi. Finally, the set of generators needs to be updated in order to obtain generators with maximal length. To establish this, we select a generator of which the new column is sufficiently dependent and replace it by the new column.

**6. Some particular cases which are also included in the analysis.** In the literature more general cases are investigated as well. In this section we discuss unions of structured spectra and draw some conclusions out of this w.r.t. the necessity of the condition of having  $H$  of a particular structured form.<sup>1</sup> Next we discuss a slightly more general class of normal matrices, admitting a relation  $s(A^*) = p(A)q(A)^{-1}$  which

---

<sup>1</sup>Proving the necessity part is far from trivial; see [2, 1] for more information.

will be used in the numerical experiments. Also it is shown that the class of matrices admitting a relation of the form  $A^* = p(A)q(A)^{-1} + R$ , with  $R$  a matrix of low rank, results in a structured rank part of the associated Hessenberg matrix. Finally, we briefly touch the class of  $B$ -normal matrices as in [22].

**6.1. Perturbations and unions of structured spectra.** In the previous paragraphs we deduced the rank structure of a normal Hessenberg of which all eigenvalues  $\lambda_i$  satisfy a relation of the form

$$(6.1) \quad \bar{\lambda}_i = p(\lambda_i)/q(\lambda_i)$$

with  $p$  and  $q$  polynomials of low degree. In this section we consider the rank structure of a normal Hessenberg for which part of the eigenvalues satisfy a relation of the form (6.1) and the remaining part of the eigenvalues satisfy another relation of the same form. This is easily generalizable to a union of more than two eigenvalue sets. We will show that there is a nonneglectable impact on the matrix structure. As an example we start with deriving the rank structure of a normal Hessenberg matrix  $H$  associated to a normal  $n \times n$  matrix  $A$  of which all but one of the eigenvalues lie on a line. We search for an orthonormal matrix  $Q$  such that  $Q^*AQ = H$  is Hessenberg. Without loss of generality we can replace  $A$  by a block matrix  $B$  of the form

$$B = \begin{bmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ & & & \times \end{bmatrix},$$

where the subblock in the upper left corner is a normal matrix of which all eigenvalues lie on a line. Choose a random starting vector  $\mathbf{v}$ . From this vector we can construct the Krylov space  $\mathcal{K}_n(B, \mathbf{v}) = \text{span}\{\mathbf{v}, B\mathbf{v}, \dots, B^{n-1}\mathbf{v}\}$ . One knows that if  $QR = K_n(B, \mathbf{v}) = [\mathbf{v}, B\mathbf{v}, \dots, B^{n-1}\mathbf{v}]$  with  $Q$  unitary and  $R$  an upper triangular invertible matrix, then  $Q^*BQ = H$  is Hessenberg. Moreover, the columns of the matrix  $Q$  satisfying this relation form an orthonormal basis of the Krylov space  $\mathcal{K}_n(B, \mathbf{v})$ . A full rank matrix has a unique  $QR$ -factorization. We will exploit this below to construct the matrix  $Q$  in a special way, which will allow us to reveal the rank structure of the Hessenberg matrix  $H$ .

It is known that one can transform a matrix to upper triangular form by a pyramid-shaped pattern of rotations. Hence, if we apply the following sequence of rotations to the matrix  $K_n(B, \mathbf{v})$ , which makes its upper left  $(n-1) \times (n-1)$  block upper triangular, we obtain

$$(6.2) \quad S = \tilde{Q}^* K_n(B, \mathbf{v}) = \begin{matrix} & & \curvearrowright & & \\ & \curvearrowright & & \curvearrowright & \\ \curvearrowright & & \curvearrowright & & \curvearrowright \\ & \curvearrowright & & \curvearrowright & \\ & & \curvearrowright & & \end{matrix} \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} = \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}.$$

Using the fact that a normal Hessenberg matrix of which all eigenvalues lie on a line is of tridiagonal form, it follows that

$$\tilde{Q}^* B \tilde{Q} = \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}.$$



Next, to remove the elements in the bottom row of (6.2) and to make  $S$  upper triangular, we apply rotations, of which each annihilates an element in the trailing row:

$$\bar{Q}^* S = \begin{array}{c} \begin{array}{c} \downarrow \\ \downarrow \\ \downarrow \\ \downarrow \end{array} \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \end{array} = \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}.$$

These rotations are accumulated in an orthonormal matrix  $\bar{Q}$ . Let  $Q = \tilde{Q}\bar{Q}$ . We know by construction that

$$Q^* B Q = \bar{Q}^* \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \\ & & & \times & \times \\ & & & & \times \end{bmatrix} \bar{Q}$$

must be of Hessenberg form. Performing this unitary similarity transformation and using the structure of  $\bar{Q}$  as a succession of rotations, it follows easily that the resulting matrix will be of structured rank form. More precisely, all subblocks above the first superdiagonal will have rank at most one, where the last column will act as a generator of this structure.

Using the same approach, one can, for example, derive the rank structure of a normal Hessenberg of which all eigenvalues but one lie on a circle. By applying rotations one finds that the last two columns of the resulting Hessenberg will generate the part of the matrix above the first subdiagonal. Hence, every submatrix lying above the first subdiagonal will have rank at most two.

As the reader already expects, one extra eigenvalue lying outside the curve containing all other eigenvalues will in general increase the low rank structure by one. This also implies that for a normal matrix having part of the eigenvalues lying on some curve and part of the eigenvalues on another curve will in general have no low rank structure, as the low rank structure grows rapidly.

To illustrate this, let  $B$  be a normal matrix with two blocks, each block capturing eigenvalues of one curve. Then there exists an orthonormal matrix  $\tilde{Q}$  such that

$$\tilde{Q}^* K(B, \mathbf{v}) = \begin{bmatrix} \tilde{Q}_1^* & \\ & \tilde{Q}_2^* \end{bmatrix} K(B, \mathbf{v}) = \begin{bmatrix} \square & \square \\ \square & \square \end{bmatrix}.$$

One has to remove now the lower left block of the latter matrix by, e.g., a sequence of rotations. One starts by a sequence of rotations  $\bar{Q}_1$  removing the first row of the lower left block, followed by a sequence of rotations  $\bar{Q}_2$  removing the second row, and so on. However, removing the lower left block will destroy the upper triangular structure of the lower right block. Hence, an extra pyramid of rotations  $\bar{Q}_{n+1}$  is needed to restore this upper triangular shape. Let  $\bar{Q} = \bar{Q}_1 \bar{Q}_2 \dots \bar{Q}_n \bar{Q}_{n+1}$  be the product of this sequence of rotations. Let  $Q = \tilde{Q}\bar{Q}$ . Then  $Q^* B Q$  will be of Hessenberg form, and the upper part will not necessary be of low rank anymore.

As an example, assume the first block of  $B$  has five eigenvalues lying on a line and the second block of  $B$  has three eigenvalues lying on a different line. Then both  $H_1$  and  $H_2$  are tridiagonal and

$$\bar{Q}_1^* \begin{bmatrix} \times & \times & & & & & \\ & \times & \times & & & & \\ & & \times & \times & \times & & \\ & & & \times & \times & & \\ & & & & \times & \times & \\ & & & & & \times & \times \\ & & & & & & \times & \times \\ & & & & & & & \times & \times \end{bmatrix} \bar{Q}_1 = \begin{bmatrix} \times & \times & \otimes & \otimes & \otimes \\ \times & \times & \times & \otimes & \otimes \\ \cdot & \times & \times & \times & \otimes \\ \cdot & \cdot & \times & \times & \otimes \\ \cdot & \cdot & \cdot & \cdot & \times & \times \\ \cdot & \cdot & \cdot & \cdot & \times & \times & \times \\ \cdot & \cdot & \cdot & \cdot & \cdot & \times & \times \end{bmatrix},$$

where the part of the matrix denoted by  $\otimes$  symbols will have a rank one structure. Note that the latter matrix can have nonzero elements below the first subdiagonal; we denote these elements by a dot. However, as  $Q^*BQ$  is Hessenberg, all elements below the first subdiagonal will become zero in the end and we do not bother about them. Next we apply  $\bar{Q}_2$  and obtain

$$\bar{Q}_2^* \bar{Q}_1^* \begin{bmatrix} \times & \times & & & & & \\ & \times & \times & & & & \\ & & \times & \times & \times & & \\ & & & \times & \times & & \\ & & & & \times & \times & \\ & & & & & \times & \times \\ & & & & & & \times & \times \\ & & & & & & & \times & \times \end{bmatrix} \bar{Q}_1 \bar{Q}_2 = \begin{bmatrix} \times & \times & \otimes & \otimes & \otimes & \otimes \\ \times & \times & \times & \otimes & \otimes & \otimes \\ \cdot & \times & \times & \times & \otimes & \otimes \\ \cdot & \cdot & \times & \times & \otimes & \otimes \\ \cdot & \cdot & \cdot & \cdot & \times & \times \\ \cdot & \cdot & \cdot & \cdot & \times & \times & \times \\ \cdot & \cdot & \cdot & \cdot & \cdot & \times & \times \end{bmatrix},$$

where now the part of the latter matrix denoted by  $\otimes$  symbols has a rank (at most) two structure. Finally, applying the last two sequences of rotations we obtain a  $7 \times 7$  Hessenberg matrix with all submatrices above the first superdiagonal having rank at most three, i.e., we cannot guarantee anything on the rank structure anymore.

**6.2. Necessary conditions for  $(\ell, m)$ -multiple recurrence relations.** In the following we discuss some difficulties in establishing necessary conditions for retrieving an  $(\ell, m)$ -multiple recurrence. For  $m = 0$  necessary conditions are found in [8, 9]; see also [1]. Recall that for a normal Hessenberg  $H$  satisfying  $H^* = p(H)q(H)^{-1}$  the degrees of the polynomials  $p(\cdot)$  and  $q(\cdot)$  induce a rank structure in the Hessenberg matrix  $H$ , resulting in a short multiple recurrence. However, the converse is not necessarily true, i.e., given an  $(\ell, m)$ -multiple recurrence, it is not necessarily derived or possible to be derived from an  $(\ell, m)$ -normal matrix. Consider, for example, a normal matrix with all eigenvalues but one on the real axis. For a generic starting vector, the corresponding Hessenberg matrix will have a rank one structure above the first superdiagonal, leading to a  $(2, 1)$ -multiple recurrence. Therefore, at first we expect this normal Hessenberg to satisfy a relation of the form  $H^* = p(H)q(H)^{-1}$  with  $p(\cdot)$  of degree 2 and  $q(\cdot)$  of degree 1. However, the distribution of the eigenvalues makes this impossible. Note that if we choose a starting vector which is in the invariant subspace spanned by all eigenvectors with eigenvalues lying on the real axis, the Hessenberg matrix belonging to the corresponding Krylov subspace of dimension  $n - 1$  will have zeros above the first superdiagonal, implying that the corresponding  $(2, 1)$ -multiple recurrence reduces to a  $(2, 0)$ -multiple recurrence. Hence, in order to establish a one-to-one correspondence between structured matrices and multiple recurrences, we have to limit ourselves to the subclass of normal matrices inducing the same low rank structure<sup>2</sup> independent of the selected starting vector. We believe

<sup>2</sup>The last column of the Hessenberg matrix is strictly speaking not of interest, although it automatically participates in the low rank structure if the eigenvalues  $\lambda$  of the Hessenberg matrix satisfy a relation of the form  $\bar{\lambda} = p(\lambda)/q(\lambda)$ .

that for this subclass of normal matrices (this also includes the matrices discussed in section 6.4 below) low rank structure of the corresponding Hessenberg  $H$  is induced by a relation of the form  $H^* = p(H)q(H)^{-1} + R$  with  $p(\cdot)$  and  $q(\cdot)$  polynomials of low degree and  $R$  a matrix of low rank. The influence of this additional low rank matrix  $R$  is discussed in section 6.4 below. We have, however, not yet been able to prove this. In this context we also refer to [1, 2], where necessary conditions were stated for  $(\ell, m)$ -multiple recurrences and were partially proved by the author, i.e., for  $(\ell, m) \leq 1$ .

**6.3. Normal matrices  $A$  of the form  $s(A^*)=p(A)q(A)^{-1}$ .** Let  $Q$  be the unitary matrix such that  $Q^*AQ = H$  is of Hessenberg form. Our aim is to retrieve this unitary matrix  $Q$  in a fast way by exploiting the relation  $s(A^*) = p(A)q(A)^{-1}$ , where  $s(\cdot)$ ,  $p(\cdot)$ , and  $q(\cdot)$  are polynomials of low degree. One can rewrite this expression as

$$Q^*s(A^*)Q = s(H^*) = p(H)q(H)^{-1}.$$

Hence, instead of computing the matrix  $H$ , one can compute the extended Hessenberg matrix  $s(H)$ , which we denote by  $\tilde{H}$ . To do this, one has to compute the first  $t$  orthonormal vectors, where  $t$  denotes the degree of  $s(\cdot)$ , using the Arnoldi recurrence

$$h_{i,i+1}\mathbf{q}_{i+1} = A\mathbf{q}_i - \sum_{j=1}^i h_{ij}\mathbf{q}_j \quad \text{with } h_{ij} = \langle A\mathbf{q}_i, \mathbf{q}_j \rangle.$$

From here on, the orthonormal vectors can be computed one by one using

$$(6.3) \quad \tilde{h}_{i,i+t}\mathbf{q}_{i+1} = s(A)\mathbf{q}_i - \sum_{j=1}^{i-1+t} \tilde{h}_{ij}\mathbf{q}_j \quad \text{with } \tilde{h}_{ij} = \langle s(A)\mathbf{q}_i, \mathbf{q}_j \rangle.$$

Of course the recurrence (6.3) is not optimal because we did not exploit the fact that  $s(H)$  has a low rank structure. More specifically, if we denote by  $\ell$  the rank of  $p(\cdot)$  and by  $m$  the rank of  $q(\cdot)$ , it has a rank  $m$  structure above the  $(\ell - m)$ th diagonal. Therefore, we can rewrite (6.3) as

$$\tilde{h}_{i,i+t}\mathbf{q}_{i+1} = \sum_{j=i-m}^i \rho_{ij}s(A)\mathbf{q}_j + \sum_{j=i-\ell}^{i-1+t} \sigma_{ij}\mathbf{q}_j$$

for some  $\rho_{ij}, \sigma_{ij}$ .

**6.4. Normal matrices of the form  $A^* = p(A)q(A)^{-1} + R$  with  $R$  a matrix of low rank.** Assume the matrix  $R$  has rank  $r$ . Again let  $Q$  be the unitary matrix such that  $Q^*AQ$  is a matrix of Hessenberg form. The relations  $A^* = p(A)q(A)^{-1} + R$  can then be rewritten as

$$H^* = p(H)q(H)^{-1} + Q^*RQ.$$

As  $Q^*RQ$  is still a matrix of rank  $r$ , the Hessenberg matrix  $H$  will have all submatrices above the  $(\ell - m)$ th diagonal of rank at most  $m + r$ . Hence, one obtains the recurrence relation

$$\mathbf{q}_{j+1} = \sum_{i=j-r-m}^j \rho_{j,i}A\mathbf{q}_i - \sum_{i=j-r-\ell}^j \sigma_{j,i}\mathbf{q}_i$$

for some  $\rho_{j,i}, \sigma_{j,i}$ .

The reader also might want to see [3], where the specific case  $A - A^* = R$ , with  $R$  a matrix of low rank, is closely investigated.

**6.5. B-normal matrices.** In [22] one considers  $B$ -normal matrices instead of normal matrices. In this paragraph we show that the theory above can be easily extended to  $B$ -normal matrices.

A matrix  $A$  is called  $B$ -normal, with  $B$  a positive definite matrix, if  $AA^\dagger = A^\dagger A$ , where  $A^\dagger = B^{-1}A^*B$ . Also, we call the matrix  $A$ ,  $B$ -normal( $\ell, m$ ) if there exists polynomials  $p$  and  $q$  of degree  $\ell$  and  $m$ , respectively, such that

$$(6.4) \quad A^\dagger = p(A)q(A)^{-1}.$$

Suppose we apply Arnoldi on the matrix  $A$  in order to retrieve a  $B$ -orthonormal basis  $Q$ . Then we obtain a Hessenberg matrix  $H = Q^{-1}AQ$  with  $Q^*BQ = I$ . Therefore,  $H^\dagger = B^{-1}H^*B = B^{-1}Q^*A^*BQB = B^{-1}Q^{-1}B^{-1}A^*BQB = (QB)^{-1}A^\dagger(QB)$ . Hence, the relation (6.4) implies  $(QB)H^\dagger(QB)^{-1} = Q(p(H)q(H)^{-1})Q^{-1}$ , or  $H^* = p(H)q(H)^{-1}$ , which is our familiar relation. Hence, we are free to replace in all theorems above the word “normal” by “ $B$ -normal,” where in the Arnoldi method we have to use the inner product  $\langle x, y \rangle = y^*Bx$ .

**7. Numerical experiments.** In this section some numerical experiments are discussed. All experiments were performed in MATLAB 7.14 (R2012a) on two hexa-core Intel Xeon E5645 CPUs with 48 GB RAM. In this paper the focus is not on the stability of the algorithm, but rather on the theoretical aspects which reveal the rank structure of a normal Hessenberg matrix and on exploiting this rank structure to obtain a short recurrence which will save us computing time with respect to the Arnoldi method. In the class of matrices  $A$  satisfying a relation of the form  $A^* = p(A)q(A)^{-1}$ ,  $p(\cdot)$ ,  $q(\cdot)$  polynomials, the two cases of practical interest are unitary and shifted unitary matrices [21]. In this context the isometric Arnoldi algorithm [14, 15, 20, 32] is introduced and compared with classical Arnoldi and our algorithm.

Furthermore, numerical results are shown for an example like the one discussed in section 6.3, revealing the numerical difficulties introduced by the use of multiple recurrences when the rank structured part consists of submatrices with a rank higher than one. One might, for instance, already wonder whether the chosen representation with generators to store the rank structured part of the Hessenberg matrix is a good numerically reliable choice [28].

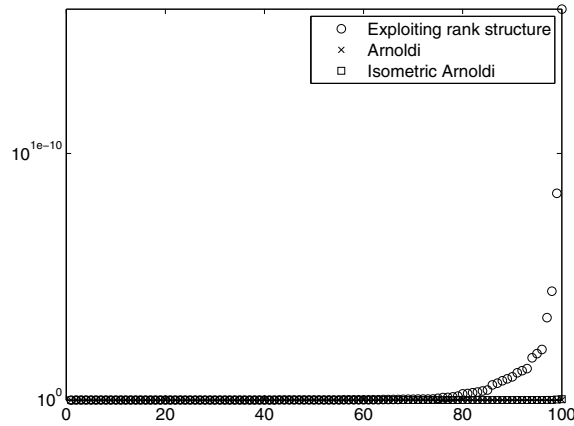
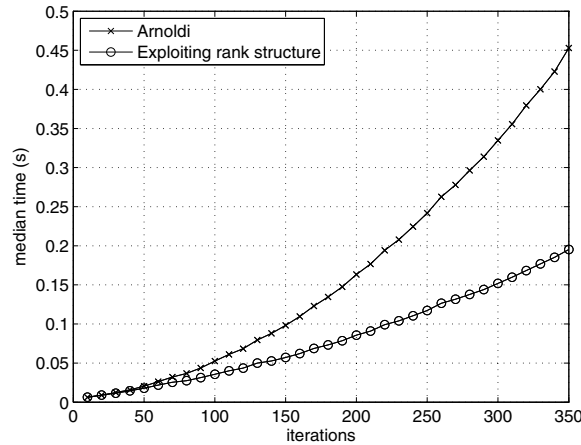
Note that in all experiments we assume to know which type of structured matrix we are dealing with, i.e., the rank structure is known, which is necessary for the application of our algorithm.

In the first experiment we consider a randomly generated unitary matrix. Let  $Q_k$  be the matrix having the computed vectors  $\mathbf{q}_i$  with  $i = 1, \dots, k$  as columns. Figure 6(a) plots the quotient of the largest singular value of  $Q_k$  divided by the smallest singular value of  $Q_k$  versus step  $k$ , for a unitary matrix of order 100. This scalar can be considered as a measure of orthogonality. We conclude that exploiting the rank structure in this case has no significant negative influence on the orthogonality.

For the specific case of unitary matrices, the isometric Arnoldi algorithm can be used. The rank one structure of the corresponding Hessenberg matrix is indirectly exploited, as the latter implies that the corresponding Hessenberg matrix is a product of rotations, giving rise to the coupled two-term recurrence relation [20]

$$\begin{aligned} \mathbf{q}_0 &= \mathbf{v}/\|\mathbf{v}\|, \quad \tilde{\mathbf{q}}_0 = \mathbf{q}_0, \\ \mathbf{q}_{j+1} &= \sigma_j^{-1}(U\mathbf{q}_j + \gamma_j\tilde{\mathbf{q}}_j), \\ \tilde{\mathbf{q}}_{j+1} &= \sigma_j\tilde{\mathbf{q}}_j + \bar{\gamma}_j\mathbf{q}_{j+1}, \end{aligned}$$

with  $\gamma_j = -\langle U\mathbf{q}_j, \tilde{\mathbf{q}}_j \rangle$ ,  $\sigma_j = ((1 - |\gamma_j|)(1 + |\gamma_j|))^{1/2}$ ,  $U$  unitary, and  $\mathbf{v} \in \mathbb{C}$ .


 (a) Measure of orthogonality of  $Q_k$ .


(b) Median execution time over 350 runs for a unitary matrix of order 1000.

FIG. 6. Numerical results on accuracy and speed for unitary matrices.

Figure 6(b) plots the average execution time for the first 350 iterations of a randomly generated unitary matrix of order 1000. Exploiting the rank structure allows us to gain a factor two in speed with respect to Arnoldi. However, in this case the isometric Arnoldi algorithm is in favor regarding complexity.

As the rank of the rank structured part grows higher, the complexity of our algorithm also increases, as each iteration step requires solving an  $m \times m$  linear system with  $m$  the rank of the rank structured part.

There are also some advantages in memory storage with respect to the Arnoldi algorithm. First of all, the storage of the Hessenberg matrix  $H$  can be made more efficient by storing the low rank part as linear combinations of the generators. Second, if the application does not demand to return the matrix  $Q$ , only part of the orthonormal vectors need to be stored. More precisely, theoretically  $m + 1 + k$  (with  $m$  the rank of the rank structured part) of the previous orthonormal vectors together with

the generators; need to be stored. The number  $k$  is variable during execution and depends on the length of the generators; the longer the generators, the smaller  $k$ . If one chooses for numerical reasons to compute  $t > m$  inner products to retrieve a new orthonormal vector,  $t + 1 + k$  of the previously computed orthonormal vectors need to be stored in order to compute the next orthonormal vector, whereas Arnoldi needs all previously computed orthonormal vectors in order to retrieve the next orthonormal vector.

As a second example we consider a normal matrix of which the eigenvalues lie on a circle whose midpoint is not located in the origin, i.e., a shifted unitary matrix  $\alpha U + \beta I$  with  $U$  unitary,  $\alpha$  and  $\beta$  constants. The eigenvalues are randomly distributed on the circle. Figure 7 shows that exploiting the rank structure does not result in a loss of orthonormality compared to modified Arnoldi. We plotted only the first 30 iterations, as after 30 iterations the orthonormality deteriorates significantly, both for our algorithm and for Arnoldi. However, if we assume that the constants  $\alpha$  and  $\beta$  are known, the isometric Arnoldi algorithm can be applied, using the observation that

$$\mathcal{K}_k(A, \mathbf{v}) = \mathcal{K}_k(U, \mathbf{v}),$$

$\mathbf{v} \in \mathbb{C}^n$ ,  $A = \alpha U + \beta I$ ,  $U$  unitary. As is shown in Figure 7, the isometric Arnoldi algorithm is able to maintain orthogonality longer. Note that for the application of our algorithm the constants  $\alpha$  and  $\beta$  should not be known.

Finally, we consider a normal matrix of order 156 of which the eigenvalues  $\lambda$  satisfy the equation  $\bar{\lambda}^2 - 1 = 1/(\lambda^2 - 1)$ . This implies that the square of the corresponding Hessenberg matrix will have a rank two structure. The algorithm of section 5 is adapted to deal with matrices from section 6.3. Figure 8(b) shows that orthonormality is preserved only during the first 20 iterations, whereas Arnoldi holds orthonormality for at least 50 iterations. Theoretically we have a structured rank part of rank two; however, numerically a strongly dominant rank one part can be seen. As a consequence, the upper structured rank part is not easily retrieved. Setting the thresholds to strict for updating generators does not update one of the generators and ends thus

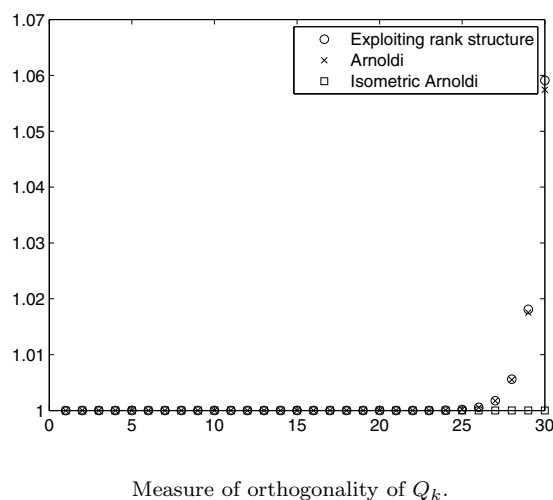
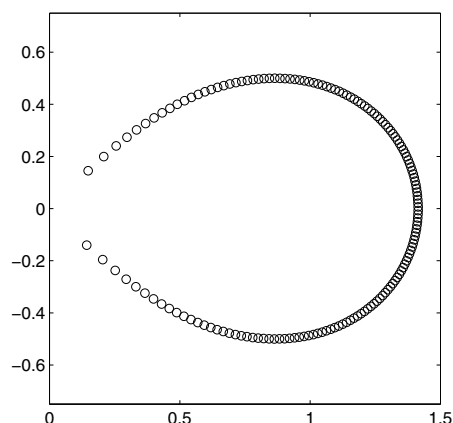


FIG. 7. Numerical results for a shifted unitary matrix of order 125.



(a) Eigenvalues of a normal matrix of order 156.

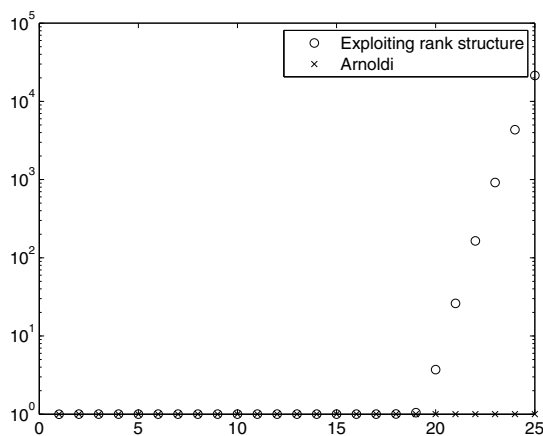

 (b) Measure of orthogonality of  $Q_k$ .

FIG. 8. Numerical results for a normal matrix of order 156 of which the square of the adjoint is a rational function of low degree in the matrix.

in a complexity comparable to Arnoldi. On the other hand, loosening the threshold for generator updating quickly results in a loss of accuracy, as seen in Figure 8(b).

**8. Conclusions.** An alternative manner of deriving multiple recurrences was presented. Instead of classically deducing the recurrences based on inner products and subspace relations, focus was on the (rank) structure of the associated Hessenberg matrix. We proved that for standard Krylov subspaces stemming from structured normal matrices, the Hessenberg matrix was of highly structured form, enabling the design of multiple recurrences.

The research proposed here imposes some natural questions. Is it possible to rely solely on the matrix structure to prove its necessity for retrieving multiple recurrences? Is it feasible to design a numerical reliable algorithm exploiting the low rank parts to save in computational resources? It is clear from our experience and the numer-



ical experiments that a straightforward implementation is certainly not sufficient to retrieve reliable orthogonal vectors. We suspect that the latter may be substantially improved by heuristically selecting a well-conditioned submatrix of the generator matrix for computing the next orthogonal vector. One such heuristic is provided by the adaptive cross-approximation algorithm [13], which attempts to select a set of rows such that the resulting submatrix approaches one of maximal volume.

**Acknowledgment.** We are grateful to both referees for their remarks and suggestions, which have helped significantly improve and clarify the manuscript. We are especially thankful for their input concerning the isometric Arnoldi algorithm and its variants for shifted unitary matrices.

## REFERENCES

- [1] T. BARTH, *Multiple Recursion Conjugate Gradient Methods*, Ph.D. thesis, Department of Mathematics, University of Colorado, Denver, 1996.
- [2] T. L. BARTH AND T. A. MANTEUFFEL, *Multiple recursion conjugate gradient algorithms, Part I: Sufficient conditions*, SIAM J. Matrix Anal. Appl., 21 (2000), pp. 768–796.
- [3] B. BECKERMANN AND L. REICHEL, *The Arnoldi process and GMRES for nearly symmetric matrices*, SIAM J. Matrix Anal. Appl., 30 (2008), pp. 102–120.
- [4] T. BELLA, Y. EIDELMAN, I. C. GOHBERG, AND V. OLSHEVSKY, *Classifications of three-term and two-term recurrence relations and digital filter structures via subclasses of quasiseparable matrices*, SIAM J. Matrix Anal. Appl., (2007), pp. 1–24.
- [5] T. BELLA, V. OLSHEVSKY, AND P. ZHLOBICH, *Classifications of recurrence relations via subclasses of  $(H, m)$ -quasiseparable matrices*, in Numerical Linear Algebra in Signals, Systems and Control, Lect. Notes Electr. Eng. 80, Springer, New York, 2011, pp. 23–53.
- [6] M. J. CANTERO, L. MORAL, AND L. VELAZQUEZ, *Five-diagonal matrices and zeros of orthogonal polynomials on the unit circle*, Linear Algebra Appl., 362 (2003), pp. 29–56.
- [7] V. FABER, J. LIESEN, AND P. TICHÝ, *The Faber-Manteuffel theorem for linear operators*, SIAM J. Numer. Anal., 46 (2008), pp. 1323–1337.
- [8] V. FABER, J. LIESEN, AND P. TICHÝ, *On orthogonal reduction to Hessenberg form with small bandwidth*, Numer. Algorithms, 51 (2009), pp. 133–142.
- [9] V. FABER AND T. MANTEUFFEL, *Necessary and sufficient conditions for the existence of a Conjugate Gradient method*, SIAM J. Numer. Anal., 21 (1984), pp. 352–362.
- [10] H. FASSBENDER AND K. D. IKRAMOV, *SYMMLQ-like procedure for  $Ax=b$  where  $A$  is a special normal matrix*, Calcolo, 43 (2006), pp. 17–37.
- [11] M. FIEDLER AND T. L. MARKHAM, *Completing a matrix when certain entries of its inverse are specified*, Linear Algebra Appl., 74 (1986), pp. 225–237.
- [12] G. H. GOLUB AND C. F. VAN LOAN, *Matrix Computations*, 3rd ed., Johns Hopkins University Press, Baltimore, MD, 1996.
- [13] S. A. GOREINOV, E. E. TYRTYSHNIKOV, AND N. L. ZAMARASHKIN, *A theory of pseudoskeleton approximations*, Linear Algebra Appl., 261 (1997), pp. 1–21.
- [14] W. B. GRAGG, *The QR algorithm for unitary Hessenberg matrices*, J. Comput. Appl. Math., 16 (1986), pp. 1–8.
- [15] W. B. GRAGG, *Positive definite Toeplitz matrices, the Arnoldi process for isometric operators, and Gaussian quadrature on the unit circle*, J. Comput. Appl. Math., 46 (1993), pp. 183–198.
- [16] R. GRONE, C. R. JOHNSON, E. M. SA, AND H. WOLKOWICZ, *Normal matrices*, Linear Algebra Appl., 87 (1987), pp. 213–225.
- [17] M. HESTENES AND E. STIEFEL, *Methods of conjugate gradients for solving linear systems*, J. Res. National Bureau of Standards, 49 (1952), pp. 409–436.
- [18] T. HUCKLE, *The Arnoldi method for normal matrices*, SIAM J. Matrix Anal. Appl., 15 (1994), pp. 479–489.
- [19] M. HUHTANEN, *Orthogonal polyanalytic polynomials and normal matrices*, Math. Comp., 72 (2002), pp. 355–372.
- [20] C. JAGELS AND L. REICHEL, *A fast minimal residual algorithm for shifted unitary matrices*, Numer. Linear Algebra Appl., 1 (1994), pp. 555–570.
- [21] J. LIESEN, *When is the adjoint of a matrix a low degree rational function in the matrix*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 1171–1180.

- [22] J. LIESEN AND Z. STRAKOŠ, *On optimal short recurrences for generating orthogonal Krylov subspace bases*, SIAM Rev., 50 (2008), pp. 485–503.
- [23] S. NOSCHESSE, L. PASQUINI, AND L. REICHEL, *The structured distance to normality of an irreducible real tridiagonal matrix*, Electron. Trans. Numer. Anal., 28 (2007), pp. 65–77.
- [24] Y. SAAD, *Numerical Methods for Large Eigenvalue Problems*, Manchester University Press, Manchester, UK, 1992.
- [25] Y. SAAD, *Iterative Methods for Sparse Linear Systems*, 2nd ed., SIAM, Philadelphia, PA, 2003.
- [26] G. STRANG AND T. NGUYEN, *The interplay of ranks of submatrices*, SIAM Rev., 46 (2004), pp. 637–646.
- [27] H. A. VAN DER VORST, *Iterative Krylov Methods for Large Linear Systems*, Cambridge Monogr. Appl. Comput. Math. 13, Cambridge University Press, Cambridge, UK, 2003.
- [28] R. VANDEBRIL, M. VAN BAREL, AND N. MASTRONARDI, *A note on the representation and definition of semiseparable matrices*, Numer. Linear Algebra Appl., 12 (2005), pp. 839–858.
- [29] R. VANDEBRIL, M. VAN BAREL, AND N. MASTRONARDI, *Matrix Computations and Semiseparable Matrices, Volume I: Linear Systems*, Johns Hopkins University Press, Baltimore, MD, 2008.
- [30] R. VANDEBRIL AND D. S. WATKINS, *A generalization of the multishift QR algorithm*, SIAM J. Matrix Anal. Appl., 33 (2012), pp. 759–779.
- [31] D. S. WATKINS, *Some perspectives on the eigenvalue problem*, SIAM Rev., 35 (1993), pp. 430–471.
- [32] D. S. WATKINS, *Unitary orthogonalization processes*, J. Comput. Appl. Math., 86 (1997), pp. 335–345.